Machine Learning	Dr. Mario Fritz and Dr. Björn Andres
Exercise Sheet 1 - 27.10.2014	due: 3.11.2014

Exercises are to be submitted in groups of three before the beginning of the lecture in lecture hall HS002. Each group submits the theoretical exercises in handwritten form, and the programming exercises via e-mail ml2014_tutors@googlegroups.com. Only one submission is required for each group. Please state your group members and assigned exercise group at the beginning of your submitted solution sheet.

Exercise 1 - Spectrum of Symmetric Matrices

Any real, symmetric matrix $A \in \mathbb{R}^{n \times n}$ has the decomposition

$$A = U \Sigma U^T,$$

where $\Sigma \in \mathbb{R}^{n \times n}$ is a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A on the diagonal and U is an orthogonal matrix in $\mathbb{R}^{n \times n}$, that is $UU^T = U^T U = \mathbb{1}$, which contains the corresponding eigenvectors (more precisely: an orthogonal basis of the eigenspace of the corresponding eigenvalue).

a. **(2 points)** Derive the eigenvalues and eigenvectors of A^k (matrix product with itself) for $k \in \mathbb{N}$.

b. **(2 points)** Prove that

$$\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \leq \lambda_{\max}(A),$$

where $\langle x, y \rangle = x^T y$ is the inner product in $\mathbb{R}^n$, $\lambda_{\max}(A)$ is the largest eigenvalue of A.

Exercise 2 - Empirical Mean and Covariance

Given a set of n points $X = [x_1, \ldots, x_n]$, where $x_n \in \mathbb{R}^d, X \in \mathbb{R}^{d \times n}$.

a. **(2 points)** Derive the minimizer c^* for function

$$f(c) = \sum_{i=1}^{n} \|x_i - c\|_2^2,$$

where $\|x\|_2$ denotes the Euclidean norm $\|c\|_2 = \sqrt{\sum_{j=1}^{d} c_j^2}$.

b. **(3 points)** Show that the empirical covariance matrix for X

$$\Sigma_X = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T$$

is positive semi-definite, that is $w^T \Sigma_X w \geq 0$, for all $w \in \mathbb{R}^d$, $\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$.
Hint: consider using the Cauchy-Schwarz inequality, $\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2$.

Exercise 3 - Multivariate Gaussian

The density function for a multivariate Gaussian $x \sim N(\mu, \Sigma)$ is defined as

$$f(x) = (2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu))$$

Now we have n multivariate Gaussian random vectors $\{x_i\}_{i=1}^n$, where $x_i \sim N(\mathbf{0}, \Sigma_i)$,$x_i \in \mathbb{R}^d$.

a. **(2 points)** Consider the case where all the random vectors are mutually independent, derive the density function for $\sum_{i=1}^{n} x_i$.

b. **(2 points)** Consider the case $n = 2$. Given the covariance matrix $cov(x_1, x_2) = C$, derive the density function for the joint vector $\binom{x_1}{x_2} \in \mathbb{R}^{2d}$.